

ZASTOSOWANIE SIECI NEURONOWYCH DO KLASYFIKACJI JAKOŚCI WLEWKÓW W PROCESIE CIĄGŁEGO ODLEWANIA STALI

WŁADYSŁAW ZALECKI

APPLICATION OF ARTIFICIAL NEURAL NETWORKS TO CLASSIFICATION OF QUALITY OF INGOTS AFTER CONTINUOUS CASTING

Abstract

The paper present primary results on an application of artificial neural networks to predictions of the influence of technological parameters of continuous casting on the quality of casted ingots. Development of the classification method for ingots was the main objective of the work. The input data composed casting parameters and quality evaluation for 852 ingots. An analysis shows that development of models and classification methods, which will predict the quality of ingots, is possible. Beta versions of statistical and neuron models were developed. It is shown in the paper that an accuracy of the models depends on the quality assessment method. Neural network classification yielded better accuracy, about 96%. An attempt of development of one general classification method led to a decrease of the quality of predictions. Classification methods require further validation. The general conclusion from the work is that at this stage separate classification methods have to be created for each steel type and for each casting machine.

1. Wprowadzenie

Z realizowanych w Instytucie Metalurgii Żelaza pracach badawczych w zakresie analizy jakości wlewków otrzymywanych w procesie ciągłego odlewania stali stosowano różnorodne metody analizy zebranych danych w tym: statystyki opisowe, metody analizy korelacji i regresji wielokrotnej oraz estymację nieliniową. Podobnie jak i w innych ośrodkach badawczych wyniki przeprowadzonych analiz nie były satysfakcjonujące, ponieważ nie doprowadziły do opracowania matematycznej zależności między parametrami technologicznymi procesu ciągłego odlewania a parametrami charakteryzującymi jakość wlewka ciągłego. Z tego względu wynikła potrzeba zbadania możliwości

wykorzystania metody sieci neuronowych do prognozowania wpływu parametrów technologicznych procesu COS na parametry opisujące jakość wlewków ciągłych. Problem ten można zaliczyć do tzw. zagadnień klasyfikacyjnych, w rozwiązywaniu których metoda sieci neuronowych może dać pozytywne wyniki.

W toku pracy dokonano przeglądu zaawansowanych metod statystycznych pod kątem ich zastosowania do rozwiązywania zagadnień z zakresu klasyfikacji. Przeanalizowano między innymi możliwość zastosowania metod: analizy dyskryminacyjnej, analizy skupień, drzew klasyfikacyjnych i skalowania wielowymiarowego.

Z przeprowadzonych analiz wstępnych wynika, że możliwe jest wykorzystanie wspomnianych metod do rozwiązywania tego typu zagadnień, jednakże bardzo

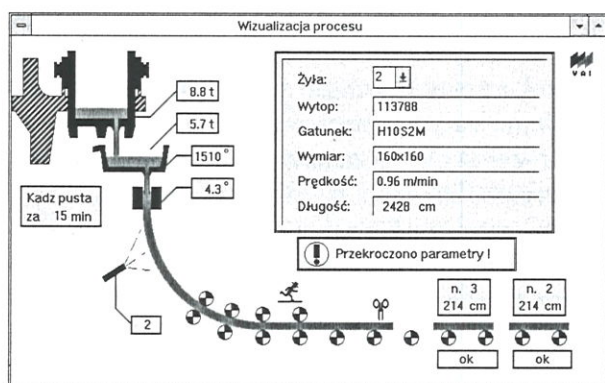
istotnego znaczenia nabiera problem zebrania i wstępnego przygotowania zbiorów danych. Stwierdzono, że w zakresie analizy jakości wlewków należy wprowadzić tzw. ogólną ocenę wlewka, ponieważ jej obecność warunkuje zastosowanie np. metody analizy dyskryminacyjnej.

W przeprowadzonych przykładowych analizach wstępnych (Zalecki i Kuziak 1999) zmienną zależną (grupującą) była ocena ogólna wlewka, a zmiennymi niezależnymi parametry oceny jakości wlewka. W takim zestawieniu zmiennych opracowany klasyfikator dokonuje klasyfikacji wlewków na podstawie wyników pomiarów parametrów opisujących ich jakość. Równie dobrze zmiennymi niezależnymi mogą być parametry technologiczne procesu COS. Uzyskane w takim przypadku funkcje klasyfikacyjne pozwolić mogą na wstępne klasyfikowanie wlewków wyłącznie na podstawie parametrów procesu, co jest celem zasadniczym i oczekiwanym przez wszystkich użytkowników maszyn COS. Takie podejście zastosowano przy poszukiwaniu modeli klasyfikacyjnych (**cel pracy**), którego przebieg przedstawiono w niniejszym referacie.

Zakres pracy obejmował:

1. Wstępne przygotowanie danych do analizy (utworzenie podzbiorów ze względu na asortyment wymiarowy i gatunkowy stali).
2. Poszukiwanie klasyfikatorów wlewków ciągłych z wykorzystaniem metody analizy dyskryminacyjnej (dla celów porównawczych).
3. Poszukiwanie klasyfikatorów wlewków ciągłych metodą sieci neuronowych.

Przedmiotem analizy był zbiór danych, zawierający parametry technologiczne procesu COS dotyczące 852 wlewków ciągłych, wraz z kryteriami oceny ogólnej jakości wlewka. W pracy wykorzystano między innymi materiały zawarte w instrukcjach zakupionych przez Instytut Metalurgii Żelaza programów Statistica® i Statistica Neural Networks®, opracowanych w firmie StatSoft® (StatSoft 1998a, StatSoft 1998b).



Rysunek 1. Schemat uproszczony urządzenia do ciągłego odlewania (wizualizacja w poziomie 2).

Ciągłe odlewanie stali jest procesem złożonym i w praktyce przemysłowej opracowanie klasyfikatorów prognozujących jakość wlewków ciągłych wymaga uwzględnienia specyfiki stosowanego urządzenia COS, asortymentu wymiarowego i gatunkowego odlewanych stali. Schemat urządzenia do ciągłego odlewania przedstawia rysunek 1.

2. WYKORZYSTANIE METODY ANALIZY DISKRYMINACYJNEJ DO POSZUKIWANIA KLASYFIKATORÓW WLEWKÓW CIĄGLYCH

Analiza funkcji dyskryminacyjnej jest jedną z metod analitycznych stosowanych do badania zbiorów wielowymiarowych. Celem analizy funkcji dyskryminacyjnej jest rozstrzygnięcie, które zmienne wyróżniają dwie lub więcej naturalnie wyłaniające się grupy.

W odniesieniu do problemów poruszanych w niniejszym referacie, metodę analizy funkcji dyskryminacyjnej można zastosować jako alternatywę lub uzupełnienie standardowej analizy statystycznej danych. Zadaniem analizy w tym przypadku było wyłonienie zmiennych, np. parametrów technologicznych procesu COS wpływających na wielkość lub liczbę wad, czyli parametrów charakteryzujących wlewkę ciągłą. W zakresie analizy procesu COS parametry procesu technologicznego oraz parametry opisujące jakość wlewka ciągłego można uznać za niezbędne do przeprowadzenia klasyfikacji. Z zebranych dotychczas doświadczeń oraz wyników obliczeń podanych w pracach (Zalacki i Kuziak 1999, StatSoft 1998a) wynika, że sformułowanie problemu klasyfikacji byłoby prostsze gdyby zebrane dotychczas dane zawierały dodatkowo ocenę ogólną wlewków np. w postaci:

1. wlewek dobry (nadaje się do dalszego przerobu),
2. wlewek warunkowo zaakceptowany (wymagane są poprawki – np. apretura powierzchni itp.) – w warunkach przemysłowych można utożsamiać ten przypadek z tzw. „wybrakiem użytkowym”,
3. wlewek zły jakościowo – wybrak.

Podstawową trudność w tym przypadku stanowić może z jednej strony wprowadzanie takiej oceny, z drugiej zaś wiarygodność przeprowadzanych ocen jakości wlewków czy półwyrobów na danym wydziale, ponieważ praktykuje się przypisywanie pewnej części wybraków innym wydziałom niż wynikałoby to ze źródła powstawania wad.

2.1 Analiza dyskryminacyjna

Podstawą analizy funkcji dyskryminacyjnej jest rozstrzygnięcie, czy grupy różnią się ze względu na średnią pewnej zmiennej, którą następnie wykorzystuje się

do przewidywania przynależności do grupy (np. nowych przypadków). Zagadnienie funkcji dyskryminacyjnej może być sprowadzone do problemu jednozmiennikowej analizy wariancji (ANOVA) (StatSoft 1998a). W szczególności, można postawić pytanie, czy dwie lub więcej grupy różnią się istotnie od siebie ze względu na średnią pewnej zmiennej (StatSoft 1998a). Jeśli średnie pewnej zmiennej są istotnie różne w poszczególnych analizowanych grupach, to możemy powiedzieć, że ta zmienna dyskryminuje (wyróżnia) te grupy (StatSoft 1998a).

W przypadku wielu zmiennych, jak ma to miejsce najczęściej (na jakość wlewków ciągłych wpływa wiele czynników), w celu sprawdzenia, które zmienne są odpowiedzialne za dyskryminowanie się grup, czy występują jakieś istotne różnice (odnośnie do wszystkich zmiennych) między grupami, porównujemy macierze (macierz całkowitych wariancji i kowariancji) wewnątrz-grupowych i międzygrupowych za pomocą wielowymiarowych testów F (StatSoft 1998a). Jest to procedura identyczna jak wielowymiarowa analiza wariancji lub MANOVA (StatSoft 1998a). Tak samo jak w wielowymiarowej analizie wariancji, możemy najpierw zastosować test wielowymiarowy, a następnie, jeśli wykazuje on istotność różnic, możemy sprawdzić, które zmienne mają istotnie różne średnie w grupach (StatSoft 1998a). Zatem chodzi o to, że szukamy zmiennych, które dyskryminują grupy, co znajduje wyraz w obserwowanych różnicach średnich (StatSoft 1998a).

Jednym z głównych celów analizy dyskryminacyjnej jest klasyfikacja analizowanych przypadków. Klasyfikacja ta odbywa się w oparciu o modele (funkcje klasyfikacyjne) opracowane na podstawie wcześniej zebranych danych – prognoza *à priori*, dając zazwyczaj gorsze wyniki niż klasyfikacja *post hoc*, gdy przewidujemy przynależność przypadków, które były użyte do estymacji funkcji dyskryminacyjnej. Jest rzeczą oczywistą, że przewidywania *post hoc* są zawsze lepsze niż przewidywania *à priori*, stąd jeśli chcemy klasyfikować przypadki przyszłe, wówczas należy zebrać nowe dane i wypróbować (ocenić) użyteczność funkcji dyskryminacyjnych (StatSoft 1998a). Dlatego otrzymane klasyfikatory należy poddać weryfikacji w oparciu o dane, które nie były użyte do ich estymacji.

Funkcje klasyfikacyjne. Moduł analizy dyskryminacyjnej programu Statistica® automatycznie oblicza funkcje klasyfikacyjne, które mogą być wykorzystane do rozstrzygnięcia do której grupy najprawdopodobniej należą poszczególne przypadki. Jest tyle funkcji klasyfikacyjnych ile grup. Każda funkcja pozwala nam obliczyć wartości klasyfikacyjne dla wszystkich przypadków z analizowanych grup, za pomocą równania (StatSoft 1998a):

$$S_i = c_i + w_{i1} * x_1 + w_{i2} * x_2 + \dots + w_{im} * x_m, \quad (1)$$

gdzie: indeks i określa daną grupę; indeksy 1, 2, ..., m określają m zmiennych; c_i jest stałą dla i -tej grupy, w_{ij} jest wagą dla j -tej zmiennej przy obliczaniu wartości klasyfikacyjnej dla i -tej grupy; x_j jest wartością obserwowaną dla danego przypadku dla j -tej zmiennej. S_i jest wypadkową wartością klasyfikacyjną (StatSoft 1998a). Funkcje klasyfikacyjne mogą być bezpośrednio wykorzystane do obliczenia wartości klasyfikacyjnych dla nowych obserwacji. Na przykład, mogą być one określone w formułach arkusza jako zależności do obliczenia nowych zmiennych; gdy do pliku zostaną dodane nowe przypadki, wartości klasyfikacyjne zostaną wówczas obliczone automatycznie (StatSoft 1998a).

Po uruchomieniu właściwego modułu w programie Statistica® i wybraniu metody analizy (tj. standardową, krokową postępującą lub wsteczną) oraz wprowadzeniu wymaganych wartości np. wartości funkcji F, stosownie do wybranej metody) otrzymujemy wyniki pozwalające na przeprowadzenie klasyfikacji przypadków nowych.

Przykład analizy

W celu znalezienia klasyfikatora jakości wlewków ciągłych analizie dyskryminacyjnej poddano zbiór (350 wytopów) zawierający 28 zmiennych (parametrów) charakteryzujących proces odlewania wlewka ciągłego płaskiego o grubości 250 mm ze stali węglowych zawierających 0,14÷0,25%C.

Jako zmienną grupującą w przeprowadzanej analizie przyjęto ogólną ocenę wytopu. Parametrami niezależnymi były zmienne charakteryzujące proces technologiczny COS. Analizę dyskryminacyjną wykonano metodą krokową postępującą. Zachowano bieżące ustawienia programowe w oknie dialogowym.

Korzystając z opcji wykresów ramkowych, na rysunku 2 pokazano przykładowe rozkłady średnich grupowych dla kilku wybranych zmiennych.

Po zakończeniu procesu obliczeniowego uzyskano model zawierający 13 zmiennych.

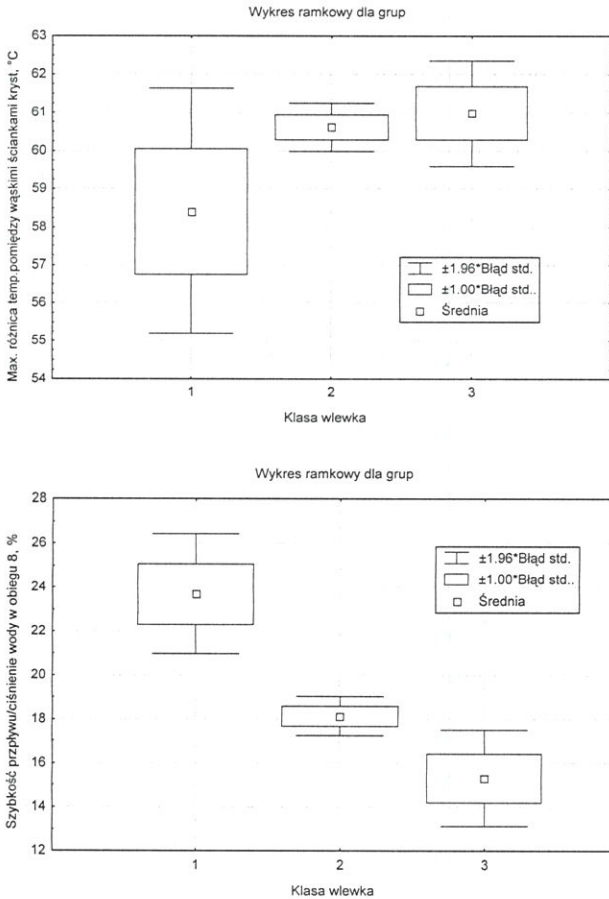
W analizie natury rozróżniania obiektu przez poszczególne zmienne może pomóc zestawienie wartości średnich zmiennych kanonicznych, które zamieszczono poniżej.

Z przedstawionych w tablicy wartości wynika, że pierwszy pierwiastek wyróżnia wybraki (wlewki złe),

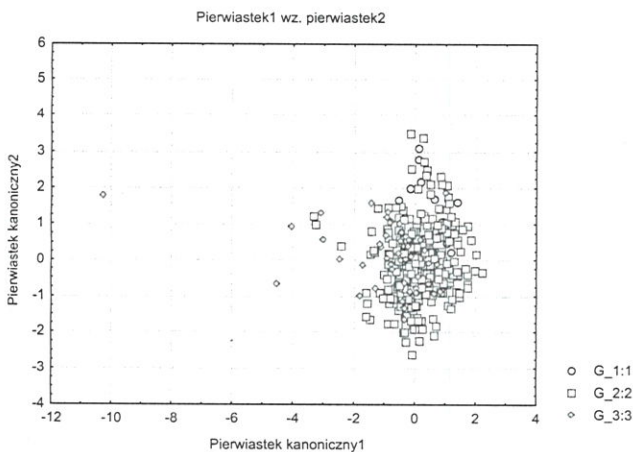
Tablica 1. Średnie zmiennych kanonicznych.

	Pierw1	Pierw2
G 1:1	0,413023	1,47279
G 2:2	0,134313	-0,07774
G 3:3	-0,89332	0,097823





Rysunek 2 Porównanie rozkładu wartości parametrów: „Maksymalna różnica temperatury pomiędzy wąskimi ściankami krystalizatora” i „Szybkość przepływu/ciśnienie wody w obiegu 8, w poszczególnych grupach charakteryzujących klasę jakości wlewka”.



Rysunek 3 Wykres punktowy natury wyróżniania poszczególnych pierwiastków kanonicznych.

drugi zaś wlewki dobre, co ilustruje wykres pokazany na rysunku 3.

Do bezpośredniego klasyfikowania analizowanych przypadków służą funkcje klasyfikacyjne, których współczynniki dla poszczególnych grup podano w tabelicy 2.

Jak wspomniano wcześniej w analizie dyskryminacyjnej można założyć różne prawdopodobieństwa a

Tablica 2. Zestawienie otrzymanych funkcji klasyfikacyjnych.

	G 1:1	G 2:2	G 3:3
	p=,03429	p=,82571	p=,14000
Odchyłka poziomu ciekłej stali w krystalizatorze	0,318584	0,316519	0,416148
Szybkość przepływu/ciśnienie w obwodzie 11	1,115041	1,219636	1,162009
Odchyłka oscylacji krystalizatora	5,58168	4,749632	5,043204
Szybkość przepływu/ciśnienie w obwodzie 10	-0,09389	-0,06029	-0,00748
Szybkość przepływu/ciśnienie w obwodzie 6	0,312542	0,295737	0,255866
Temperatura powierzchni wlewka ciągłego	-0,00536	-0,00518	-0,0062
Szybkość przepływu/ciśnienie w obwodzie 9	-0,56351	-0,69963	-0,69402
Szybkość przepływu/ciśnienie w obwodzie 7	0,766155	0,585372	0,628195
Szybkość przepływu/ciśnienie w obwodzie 1	0,285989	0,350163	0,319941
Szybkość przepływu/ciśnienie w obwodzie 4	0,240618	0,172776	0,123744
Szybkość przepływu/ciśnienie w obwodzie 8	0,807179	0,804686	0,747572
Szybkość przepływu/ciśnienie w obwodzie 13	-0,21074	-0,12279	-0,08731
Ciśnienie gazu każd. gł./każ. pośrednia	4,180352	4,289343	4,294917
Stała	-206,439	-207,602	-209,307

priori dla każdej z grup, zależnie od znajomości analizowanego obiektu. W naszej analizie przyjęto bieżące ustawienia programowe (proporcjonalne do wiel-

Tablica 3. Macierz klasyfikacji.

	Procent Poprawne	G_1:1	G_2:2	G_3:3
		p=,03429	p=,82571	p=,14000
G_1:1	16,67	2	10	0
G_2:2	98,27	2	284	3
G_3:3	12,25	0	43	6
Razem	83,43	4	337	9

kości grup). Otrzymałą macierz klasyfikacji pokazano w tabelicy 3.

W pierwszej kolumnie pokazano procent poprawnie sklasyfikowanych przypadków, a w pozostałych udział przypadków błędnie sklasyfikowanych w poszczególnych grupach.

Na 350 analizowanych wlewków 12 wlewków należało do klasy 1, 289 stanowiły wlewki (wybraki użytkowe) w klasie 2 i pozostałe 49 to wybraki (klasa 3). Ogólna trafność opracowanego modelu wynosi 83,43% ($100\% \cdot (2+284+6)/350$). Z przytoczonej w tabelicy 3 macierzy klasyfikacji można zauważyć, że separacja klasy 1 i 3 jest bardzo dobra, tzn. żaden z wlewków w kl.1 nie został zakwalifikowany przez opracowany wstępny model do klasy 3 i na odwrót.

Jest to ważne stwierdzenie ponieważ przy praktycznym wykorzystaniu opracowanych modeli największe niebezpieczeństwo błędnej klasyfikacji stanowi sklasyfikowanie wlewka wadliwego do grupy wlewków dobrych (klasa 1). Osobną uwagę należy zwrócić na klasyfikowanie wlewków złych (klasa 3) i tzw. wybraków użytkowych (lub inaczej wlewków, których dalsze zastosowanie wymaga szczegółowego zbadania i dokonania przeklasyfikowania na inne wyroby lub naprawy). Mimo stosunkowo gorszej zdolności rozdzielania wlewków klas 2 i 3 dla zastosowań praktycznych nie miałyby to większego znaczenia, gdyby nie fakt, że najliczniejszą grupę stanowią właśnie wlewki klasy 2. Przy postępującej optymalizacji procesu COS należy dążyć do zwiększenia udziału wlewków klasy 1, kosztem wlewków klasy 2 i zmniejszając udział wlewków klasy 3.

Przytoczony przykład analizy wskazuje na możliwość zastosowania metody do klasyfikowania wlewków ciągłych pod warunkiem dysponowania ogólną oceną wlewka. Uzyskane w przeprowadzonej analizie funkcje klasyfikacyjne pozwolić mogą na wstępne klasyfikowanie wlewków wyłącznie na podstawie parametrów procesu.

Uzyskany w toku analizy model klasyfikacji wlewków musi być poddany weryfikacji. Należy przy tym oczekiwać pogorszenia zdolności klasyfikacyjnych tym większego, im większa jest różnica przedziałów zmienności poszczególnych parametrów procesu COS w przypadku wlewków analizowanych przy tworzeniu modelu i zawartych w zbiorze weryfikacyjnym.

3. ZASTOSOWANIE SIECI NEURONOWYCH DO ROZWIĄZYWANIA ZAGADNIENIA KLASYFIKACJI

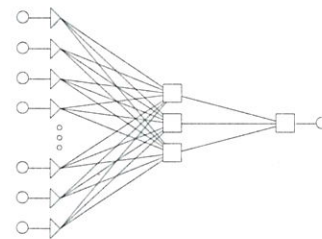
Klasyfikacja jest najbardziej rozpowszechnionym problemem rozwiązywanym za pomocą sieci neuronowych. Zadaniem sieci jest przypisanie (sklasyfikowanie) każdego z analizowanych przypadków do jednej lub kilku klas (czyli oszacowanie prawdopodobieństwa przynależności danego przypadku do każdej z klas) (StatSoft 1998b). W tym też celu stosowane są sieci o różnej architekturze jak np. sieć wielowarstwowa jednokierunkowa (Masters 1996). W programie Statistica Neural Networks™, problem klasyfikacji można rozwiązać stosując następujące typy sieci: MLP (perceptron wielowarstwowy), RBF, Kohonena, PNN i liniowej (StatSoft 1998b).

W niniejszej pracy do budowy neuronowego modelu klasyfikacyjnego wykorzystano opcję automatycznego projektanta sieci, który poszukiwał architektury sieci optymalnej dla postawionego zadania.

Zastosowano różne dane wejściowe, jak również różny podział zbioru na podzbiory uczący, weryfikujący i testowy.

3.1. Neuronowy model klasyfikacyjny wlewków – sieć uczona wyłącznie na danych liczbowych

W pierwszej kolejności przeprowadzono próbę zbudowania modelu klasyfikującego wlewki ciągle z wykorzystaniem sieci neuronowej, której procesy tworzenia i uczenia zostały przeprowadzone z wykorzystaniem zbioru danych numerycznych (350 wytopów). Zmiennymi wejściowymi były parametry technologiczne procesu COS, zmienną wyjściową była klasa wlewka oznaczona liczbą z przedziału 1÷3.



Rysunek 4. Architektura sieci klasyfikującej wlewki odlane w procesie COS (format wlewka 250 mm, grupa stali węglowych zawierających 0,14÷0,25%C).

W wyniku przeprowadzonych obliczeń wygenerowana została sieć typu MLP (perceptron wielowarstwowy), której architekturę przedstawiono na rysunku 4.

Z rysunku 4 wynika, że otrzymana sieć ma jeden neuron wyjściowy. W pierwszej z 3 warstw znajduje się 28 neuronów wejściowych o liniowej funkcji aktywacji i liniowej funkcji PSP (post synaptic potential). W drugiej warstwie mamy trzy neurony o logistycznej funkcji aktywacji i liniowej funkcji PSP. Ostatnią warstwę tworzy jeden neuron wyjściowy o logistycznej funkcji aktywacji i liniowej funkcji PSP.

Proces uczenia przeprowadzony został automatycznie, przy ustawieniu następującego podziału losowego zbioru danych:

1. zbiór uczący: 280 wytopów,
2. zbiór weryfikujący: 50 wytopów,
3. zbiór testujący: 20 wytopów.

Po wytrenowaniu na wejście sieci podano pełny zbiór danych w celu określenia jej zdolności klasyfikacyjnych. Uzyskana trafność klasyfikacji wlewków wyniosła 86,57%.

W następnej kolejności uczeniu poddano sieć o identycznej architekturze, ale cały proces przeprowadzono na pełnym zbiorze danych (350 wlewków). Sieć



podano uczeniu metodą sprzężonego gradientu przy ustawieniu 5000 epok. Uzyskana trafność klasyfikacji wlewków w tym przypadku wyniosła 95,14%.

W tym miejscu należy nadmienić, że uzyskana poprawa zdolności klasyfikacyjnych może okazać się w praktyce pozorna, ponieważ sieć była uczona na tym samym zbiorze. Zatem w tym przypadku szczególnie znaczenia nabiera weryfikacja tego modelu.

Zbudowane modele, które omówiono wcześniej mają jedną niedogodność. Na wyjściu sieci neuronowej pojawia się wartość, która mieści się w pewnym otoczeniu liczb 1,2 i 3 i stąd przy przenoszeniu wyników obliczeń do np. arkusza kalkulacyjnego odpowiedzi sieci należy zaokrąglić do najbliższej liczby całkowitej.

3.2. Neuronowy model klasyfikacyjny wlewków – wyjście sieci zmienna tekstowa

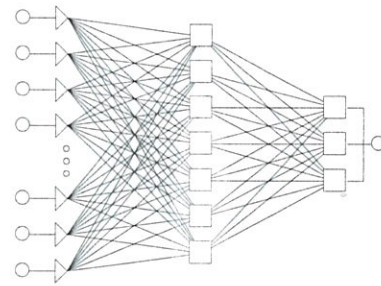
Niedogodność wspomniana na zakończenie poprzedniego rozdziału może być usunięta, poprzez zdefiniowanie zmiennej wyjściowej jako typ nominalny podając odpowiednie wartości np. dobry, wybrak_u i wybrak.

W pierwszej kolejności przeprowadzono próbę zbudowania modelu klasyfikującego wlewkę ciągłą z wykorzystaniem sieci neuronowej, której procesy tworzenia i uczenia zostały przeprowadzone z wykorzystaniem zbioru danych podzielonego w następujący sposób:

Proces uczenia przeprowadzony został automatycznie, przy ustawieniu następującego podziału losowego zbioru danych:

1. zbiór uczący: 280 wytopów,
2. zbiór weryfikujący: 50 wytopów,
3. zbiór testujący: 20 wytopów.

W wyniku przeprowadzonych obliczeń wygenerowana została sieć typu MLP (perceptron wielowarstwowy), której architekturę przedstawiono na rysunku 5.



Rysunek 5. Architektura sieci klasyfikującej wlewkę odlaną w procesie COS (format wlewka 250 mm, grupa stali węglowych zawierających 0,14÷0,25%C).

Z rysunku 5 wynika, że otrzymana sieć ma trzy neurony wyjściowe. W pierwszej z 3 warstw znajduje się 28 neuronów wejściowych o liniowej funkcji aktywacji i liniowej funkcji PSP (post synaptic potential). W drugiej warstwie mamy 7 neuronów o logistycznej funkcji aktywacji i liniowej funkcji PSP. Ostatnią warstwę tworzą trzy neurony o logistycznej funkcji aktywacji i liniowej funkcji PSP.

Po wytrenowaniu sieci zbadano jej zdolności klasyfikacyjne. Okazało się, że zdolność klasyfikacji zależy od ustawień tzw. postprocesora, a ściślej od wybrania progów akceptacji i odrzucenia. Masters (1996) uważa, że sieć w problemach klasyfikacji musi uczyć się wytwarzać realizowalny poziom aktywacji, a nie wartości ekstremalne, które są trudne lub niemożliwe do wytworzenia. Na przykład w sieciach o logistycznej funkcji aktywacji nigdy nie uzyska się pełnego poziomu aktywacji o wartościach 1,0 lub 0,0. Jeżeli będziemy starać się zmusić sieć do uzyskania takich wartości to może pojawić się niestabilność numeryczna związana z ograniczeniem zakresu liczbowego (wagi będą podciągane do bardzo dużych wartości). Według Mastersa (1996) najlepiej uczyć sieć osiągania umiarkowanych wartości. A zatem w przypadku użytej w naszej sieci – logistycznej funkcji aktywacji tradycyjnie wybiera się wartość 0,1 dla jednej decyzji i 0,9 dla drugiej.

Najważniejszym wskaźnikiem zdolności klasyfikacyjnych sieci neuronowej w programie Statistica Neu-

Tablica 4. Wyniki klasyfikacji wlewków: próg akceptacji 0,95; próg odrzucenia 0,05.

0,95/0,05	Zbiór uczący			Zbiór weryfikujący			Zbiór testowy		
	dobry	wybrak_u	wybrak	dobry	wybrak_u	wybrak	dobry	wybrak_u	wybrak
Razem	10	228	42	2	42	6	0	19	1
Poprawn.	0	226	0	0	33	0	0	15	0
Źle	3	0	15	2	0	4	0	0	0
Niesklas.	7	2	27	0	9	2	0	4	1
dobry	0	0	0	0	0	0	0	0	0
wybrak_u	3	226	15	2	33	4	0	15	0
wybrak	0	0	0	0	0	0	0	0	0

Tablica 5. Wyniki klasyfikacji wlewków próg akceptacji 0,9; próg odrzucenia 0,1.

0,9/0,1	Zbiór uczący			Zbiór weryfikujący			Zbiór testowy		
	dobry	wybrak_u	wybrak	dobry	wybrak_u	wybrak	dobry	wybrak_u	wybrak
Razem	10	228	42	2	42	6	0	19	1
Poprawn.	0	226	0	0	33	0	0	15	0
Źle	3	0	15	2	0	4	0	0	0
Niesklas.	7	2	27	0	9	2	0	4	1
dobry	0	0	0	0	0	0	0	0	0
wybrak_u	3	226	15	2	33	4	0	15	0
wybrak	0	0	0	0	0	0	0	0	0

Tablica 6. Wyniki klasyfikacji wlewków próg akceptacji 0,7; próg odrzucenia 0,3.

0,7/0,3	Zbiór uczący			Zbiór weryfikujący			Zbiór testowy		
	dobry	wybrak_u	wybrak	dobry	wybrak_u	wybrak	dobry	wybrak_u	wybrak
Razem	10	228	42	2	42	6	0	19	1
Poprawn.	0	226	0	0	35	0	0	15	0
Źle	3	0	15	2	0	5	0	0	0
Niesklas.	7	2	27	0	7	1	0	4	1
dobry	0	0	0	0	0	0	0	0	0
wybrak_u	3	226	15	2	35	5	0	15	0
wybrak	0	0	0	0	0	0	0	0	0

ral Networks™ (StatSoft 1998b) jest tzw. „okno statystyki klasyfikacyjnej”. Pokazuje ono ile przypadków zostało poprawnie sklasyfikowanych, ile niepoprawnie, a ile nie zostało sklasyfikowanych w ogóle. „Okno statystyki klasyfikacyjnej” pozwala również stwierdzić ile przypadków należących do każdej z klas zostało przypisanych do innej klasy. Wszystkie te wielkości są podawane niezależnie dla danych w fazach: uczenia, weryfikacji i testowania (StatSoft 1998b). W celu przesłedzenia trafności opracowywanych klasyfikatorów neuronowych przyjęto programowe ustawienia progów akceptacji i odrzucenia (odpowiednio 0,95 i 0,05). Uzyskane wyniki pokazano w tablicy 4.

Z analizy tablicy 4 wynika, że sieć nie rozpoznaje wlewków. W kolejnych krokach zmieniano wartości progów akceptacji i odrzucenia, a wyniki zestawiono w tablicach 5 i 6.

Z porównania tablic 4÷6 wynika, że zmieniając ustawienia progów akceptacji i odrzucenia można poprawić zdolność klasyfikującą sieci z 78,29% (tablica 4 do ok. 78,85% tablica 6). Otrzymane wyniki wskazują, że sieć potrafi identyfikować wyłącznie wlewiki klasy 2 (wybrak użytkowy), które stanowiły najliczniejszą grupę. Dlatego postanowiono powtórzyć proces uczenia na połączonym zbiorze. Zatem uczeniu poddano sieć o identycznej architekturze, ale cały proces przeprowadzono na pełnym zbiorze danych (tj. 350 wlewków). Proces uczenia przeprowadzono metodą gradientu sprzężonego opadającego (conjugate gradient descent) do uzyskania minimum (prawdopodob-

Tablica 7. Wyniki klasyfikacji wlewków: próg akceptacji 0,95; próg odrzucenia 0,05.

0,95/0,05	dobry	wybrak_u	wybrak
Razem	12	289	49
Poprawn.	5	286	43
Źle	4	0	4
Niesklas.	3	3	2
Dobry	5	0	0
Wybrak_u	4	286	4
Wybrak	0	0	43

Tablica 8. Wyniki klasyfikacji wlewków próg akceptacji 0,9; próg odrzucenia 0,1.

0,9/0,1	dobry	wybrak_u	wybrak
Razem	12	289	49
Poprawn.	7	286	43
Źle	4	0	4
Niesklas.	1	3	2
Dobry	7	0	0
Wybrak_u	4	286	4
Wybrak	0	0	43

nie lokalnego), co nastąpiło po ok. 2200 epokach.

Z analizy wyników zamieszczonych w tablicach 7÷9 wynika, że trafność klasyfikacji wlewków w sieci nauczanej na połączonym zbiorze wzrasta z ok. 95,43% (tablica 7) do ok. 96,0% (tablica 9), przy czym wystarcza już ustawienie wartości progu akceptacji 0,9 i progu odrzucenia 0,1 (tablica 8).



Tablica 9. Wyniki klasyfikacji wlewków próg akceptacji 0,7; próg odrzucenia 0,3.

0,7/0,3	dobry	wybrak_u	wybrak
Razem	12	289	49
Poprawn.	7	286	43
Źle	4	0	4
Niesklas.	1	3	2
dobry	7	0	0
Wybrak_u	4	286	4
Wybrak	0	0	43

Również w tym samym miejscu należy dodać, że uzyskana poprawa zdolności klasyfikacyjnych może okazać się pozorną, ponieważ sieć była uczona na tym samym zbiorze. Zatem i w tym przypadku szczególnego znaczenia nabiera weryfikacja opracowanego modelu.

W trakcie realizacji pracy przeprowadzono próby zbudowania klasyfikatorów na całym zbiorze danych (bez podziału na asortyment wlewków i asortyment gatunkowy odlewanej stali), ale uzyskane wyniki (nie zamieszczone w niniejszym opracowaniu) okazały się znacznie gorsze – dla sieci z wyjściem numerycznym (zbiór uczący 852 wytopy) trafność prognozowania wynosiła ok. 76%.

3.3. Analiza wpływu rodzaju funkcji aktywacji i błędu na trafność prognozowania klasyfikatora neuronowego

Instrukcja programu Statistica Neural Network™ (StatSoft 1998b) w rozdziale poświęconym rozwiązywaniu zagadnień klasyfikacyjnym podaje, że poprawę trafności prognozowania klasyfikatora neuronowego można uzyskać zmieniając rodzaje funkcji aktywacji i błędu, względem którego prowadzony jest proces uczenia sieci.

Wyjście sieci neuronowych realizujące zagadnienie klasyfikacji złożonej z więcej niż dwóch klas może być interpretowane jako prawdopodobieństwa jeżeli wartości wyjścia będą mieścić się w przedziale 0÷1. Aby to uzyskać jako funkcję aktywacji neuronów w

Tablica 10. Wyniki klasyfikacji wlewków próg akceptacji 0,9; próg odrzucenia 0,1.

0,9/0,1	dobry	wybrak_u	wybrak
Razem	12	289	49
Poprawn.	0	0	0
Źle	0	0	0
Niesklas.	12	289	49
Dobry	0	0	0
Wybrak_u	0	0	0
Wybrak	0	0	0

Tablica 11. Wyniki klasyfikacji wlewków próg akceptacji 0,7; próg odrzucenia 0,3.

0,7/0,3	dobry	wybrak_u	wybrak
Razem	12	289	49
Poprawn.	0	16	0
Źle	0	0	5
Niesklas.	12	273	44
Dobry	0	0	0
Wybrak_u	0	16	5
Wybrak	0	0	0

Tablica 12. Wyniki klasyfikacji wlewków próg akceptacji 0,65; próg odrzucenia 0,35.

0,65/0,35	dobry	wybrak_u	wybrak
Razem	12	289	49
Poprawn.	0	102	0
Źle	8	0	11
Niesklas.	4	187	38
Dobry	0	0	0
Wybrak_u	8	102	11
Wybrak	0	0	0

warstwie wyjściowej za pomocą edytora sieci należy wybrać funkcję „Softmax” i równocześnie funkcję błędu ustawić jako „Entropy (multiple)” (StatSoft 1998b).

Obliczenia przeprowadzono na zbiorze połączonym tj. zawierającym 350 wlewków. Wyniki dla różnych wartości progów akceptacji i odrzucenia przedstawiono w tablicach 10÷12. Okazuje się, że w tym przypadku zmiana funkcji aktywacji i błędu nie przyniosła poprawy zdolności prognostycznych klasyfikatora neuronowego. Dopiero przy ustawieniu wartości progu akceptacji 0,65 i progu odrzucenia 0,35 sieć rozpoznała poprawnie jakość 102 z 350 wlewków.

4. PODSUMOWANIE

Przedstawione w artykule wyniki analiz wskazują, że możliwe jest zbudowanie modeli, a na ich podstawie klasyfikatorów wlewków ciągłych, pozwalających prognozować jakość wlewków odlanych w procesie COS. Trafność wstępnie opracowanych modeli zależy przede wszystkim od przyjętego sposobu oceny wlewków i od zastosowanej metody, przy czym w przypadku klasyfikatorów neuronowych uzyskano lepsze wyniki (ok. 96%). W pracy przeprowadzono próby zbudowania klasyfikatorów dla „globalnego” zestawu danych (wyniki tych prób nie zamieszczono), jednakże trafność przewidywania jakości wlewków była znacząco niższa.

Zbudowane klasyfikatory wymagają dalszej weryfikacji. Z przeprowadzonej analizy wynika również, że tego rodzaju klasyfikatory muszą być tworzone

osobno dla każdego rodzaju (asortymentu) wlewka, grupy gatunkowej stali i maszyny COS.

Wyróżnikiem obu zastosowanych metod jest między innymi fakt, że w modelu statystycznym liczba zmiennych niezależnych (parametrów technologicznych) ulega zmniejszeniu zgodnie z przeprowadzoną weryfikacją istotności wpływu danego czynnika. W przypadku sieci neuronowej wszystkie czynniki pozostają w modelu, a ich praktyczna istotność wpływu odzwierciedlona jest wartościami wag.

Wyniki uzyskane w niniejszej pracy pozwalają na wysunięcie wniosków pozwalających ukierunkować prowadzone prace w zakresie prognozowania jakości wlewków ciągłych.

1. Konieczne jest w gromadzonych bazach danych rejestrowanie tzw. ogólnej oceny jakości wlewków.
2. Zbierane parametry procesu COS muszą odnosić

się do ocenionego pod względem jakości odcinka wlewka.

Dlatego należy równolegle gromadzić dane w zakresie pozwalającym określić jakość globalną wlewka na podstawie ocen cząstkowych, oraz parametry procesu COS.

LITERATURA

- Zalecki W., Kuziak R., 1999, Zastosowanie sieci neuronowych do analizy jakości wlewków w procesie ciągłego odlewania stali, *Sprawozdanie IMŻ nr S-00268/1/BM/99* (nie publikowane).
- StatSoft®, 1998a, Statistica - Volume III (Statistics II) - instrukcja programu
- StatSoft®, 1998b, Statistica Neural Networks - instrukcja programu.
- Masters T., 1996, *Sieci neuronowe w praktyce*, WNT, Warszawa.

